



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
-----------------	-------------	----------------------	---------------------	------------------

10/762,060

01/21/2004

Min Chu

M61.12-0594

4639

27366

7590

06/02/2009

WESTMAN CHAMPLIN (MICROSOFT CORPORATION)

SUITE 1400

900 SECOND AVENUE SOUTH

MINNEAPOLIS, MN 55402

EXAMINER

SHAH, PARAS D

ART UNIT

PAPER NUMBER

2626

MAIL DATE

DELIVERY MODE

06/02/2009

PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Office Action Summary	Application No. 10/762,060	Applicant(s) CHU ET AL.	
	Examiner PARAS SHAH	Art Unit 2626	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 13 March 2009.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1,5,6,9,12-14,17-20,22-24 and 29 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 1, 5, 6, 9, 12-14, 17-20, 22-24, 29 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|--|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413)
Paper No(s)/Mail Date. _____ |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | 5) <input type="checkbox"/> Notice of Informal Patent Application |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

1. This communication is in response to the Amendments and Arguments filed on 03/13/2009. Claims 1, 5, 6, 9, 12-14, 17-20, 22-24, and 29 are pending and have been examined with claims 4, 21, and 33 being cancelled. The Applicants' amendment and remarks have been carefully considered, but they are not persuasive and do not place the claims in condition for allowance. Accordingly, this action has been made FINAL.
2. All previous objections and rejections directed to the Applicant's disclosure and claims not discussed in this Office Action have been withdrawn by the Examiner.

Response to Arguments

3. Applicant's arguments (pages 8-14) filed on 03/13/2009 with regard to claims 1-29 have been fully considered but they are not persuasive for the reasons mentioned below.
4. With respect to independent claims 1, 20, and 29, the Applicants argue that the reference of Chen does not render obvious the claimed invention of a categorical level of pitch is assigned to each of the temporal portions. However, the Examiner respectfully disagrees. In Chen, Figure 3, and col. 4, lines 12-13 and lines 33-35, the five tones described consist of a pitch contour, which varies with respect to time for the syllable. Thus, there is a specific pitch that is measurable with respect to time, thus indicating the presence of a categorical level for a specific time instance. Further, the Applicant's argue that the model of the final part is broken time-wise into a first portion and a second portion. However, it should be noted that such breaking time-wise is not

present in the claimed limitations, where the claims recite that the final portion comprises a first and second portion and further argue that the teachings of the present disclosure were relied upon rather than the teachings of the prior art. Chen does teach such a structure based on the beginning of the syllable for a rising tone, of a part of a syllable, being a low value being raised to a higher pitch value as time progresses (e.g. near the end of the part) (see Figure 3 and col. 4, lines 12-13, and lines 33-35).

Furthermore, the Applicants argue that Chen does not teach “assign to each of those portion a discrete categorical level of pitch.” In response to applicant's argument that the references fail to show certain features of applicant's invention, it is noted that the features upon which applicant relies (i.e., “an unchanging, constant, discrete categorical level of pitch”) are not recited in the rejected claim(s). Although the claims are interpreted in light of the specification, limitations from the specification are not read into the claims. See *In re Van Geuns*, 988 F.2d 1181, 26 USPQ2d 1057 (Fed. Cir. 1993).

As to claims 6, 14, and 33, the Applicants argue that the categorical levels are not analogous to Chen's five tone types. The Examiner respectfully disagrees. Chen's five tone types denote the pitch contour for each tone. The pitch contour represent a time-wise representation of the pitch for the part of the syllable. Further, it should be noted that the claim 1, from which it depends, does not distinguish that each tone has different levels of pitch, but rather indicates that the different tones have different levels of pitch, which is broad enough to read on Chen, where Chen has 5 tones (e.g. different tones) and 5 pitch contours (different pitch contours) and they contain five categorical levels (high, rising, falling), where the association is implied by the pitch contour.

The rejections with respect to the dependent claims not mentioned above, are similarly rejected for the arguments presented above.

Claim Rejections - 35 USC § 103

5. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

6. Claims 1, 6, 9, 12, 14, 17, and 29 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chen *et al.* (US 5,751,905) in view of Huang *et al.* ("Whistler: A trainable Text-to-Speech System", 1996).

As to claims 1, 9, and 29, Chen *et al.* teaches

a speech processing system receiving an input related to one of speech and process the input to provide an output related to one of text (see Figure 6, input into microphone 600, the output of related information would have been obvious to Chen as the system is for use in speech recognition), the speech processing system (see col. 6, lines 26-36) accessing a module (see col. 3, lines 61-col. 4, lines 8, observations used within the toned phoneme system) derived from a phone set having a plurality of phones for a tonal language (see col. 4, lines 41-44, initials with glides and a second part (final)), wherein the tonal language comprises a plurality of different tones with different levels of pitch (see col. 4, lines 31-35, each tone has an associated pitch contour) the phones being

used to model syllables used in the module (see col. 6, lines 42-45), the syllables having an initial part and final part (see col. 6, lines 42-45), wherein at least some of the syllables of the tonal language include a glide, the glide being embodied in the initial part (see col. 4, lines 42-43, glide is grouped with the initial) and wherein the final part comprises a first portion corresponding to a first relative pitch and a second portion corresponding to a second relative pitch, wherein the first portion and the second portion jointly and implicitly carry the tonal information (see col. 4, lines 10-13 and col. 4, lines 42-45, the pitch contour varies with time so the pitch changes relative to the portion of the phone i.e. if the phoneme is associated with a rising pitch contour, such a contour is representing a pitch increasing from a base value); and wherein the different levels of pitch comprises at least two categorical levels (see col. 4, lines 33-35, five types of tones), and wherein each portion has a categorical level associated with it (see col. 4, lines 10-15, pitch varies with time and represents a pitch contour. The contour consist of different level or values with respect to time) (E.g. Hence, as the pitch varies over the duration of the syllable the use of categorical levels for each portion vary based on the identified tone. For example, rising tone goes from a low to high value (two categorical levels)).

However, Chen *et al.* does not specifically teach the input being text and the output being speech.

Huang *et al.* does teach the conversion of text to speech from learning methods of model parameters (see Abstract).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech processing system taught by Chen et al. and include a text to speech converter taught by Huang et al. The motivation to have included such an element is to have an alternative means for inputting as well as producing a synthesized speech output based upon model parameters of the system (see Huang et al., Abstract) as would benefit the system of Chen et al. by using the tone related information as output speech for producing speech resembling the user.

As to claim 29, Chen in view of Huang teach all of the limitations as in claim 1, above and further teach the computer readable storage medium (see col. 8, lines 28, multipurpose computer.) The use of a computer readable storage medium is obvious to one skilled in the art.

As to claims 12, Chen et al. in view of Huang et al. teaches all of the limitations as in claim 1, above.

Furthermore, Chen et al. teaches wherein the different levels of pitch comprises two categorical levels (see col. 4, lines 33-35, five types of tones), and wherein each portion has a categorical level associated with it (see col. 4, lines 10-15, pitch varies with time and represents a pitch contour. The contour consist of different level or values with respect to time) (E.g. Hence, as the pitch varies over the duration of the syllable the use of categorical levels for each portion

Art Unit: 2626

vary based on the identified tone. For example, rising tone goes from a low to high value (two categorical levels)).

As to claim 6, 14, and 33 Chen *et al.* in view of Huang *et al.* teaches all of the limitations as in claim 1, above.

Furthermore, Chen *et al.* teaches wherein the different levels of pitch comprises five categorical levels (see col. 4, lines 33-35, five types of tones), and wherein each portion has a categorical level associated with it (see col. 4, lines 10-15, pitch varies with time and represents a pitch contour. The contour consist of different level or values with respect to time).

As to claim 17, Chen *et al.* in view of Huang *et al.* teaches all of the limitations as in claim 1, above.

Furthermore, Chen *et al.* teaches wherein the tonal language comprises Chinese or a dialect thereof, such as Cantonese (see coll. 3, lines 63-64, Mandarin Chinese).

7. Claims 5 and 13 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chen in view of Huang *et al.* as applied to claims 1 and 9 above, and further in view of Akinlabi *et al.* ("tonal Phonology of Yoruba Clitics").

As to claims 5 and 13, Chen in view of Huang *et al.* teach the phone being associated with a categorical level and the limitations as in claims 1 and 9, above.

However, they do not specifically teach the levels of pitch comprising three categorical levels.

Akinlabi *et al.* teaches three types of tones being associated phonemically (see page 2, sect. 2, lines 1-2).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech processing system taught by Chen *et al.* in view of Huang *et al.* with three categorical levels taught by Akinlabi *et al.* The motivation to have included five categorical levels involves the inclusion of other tone languages such as Yoruba, where three tones are present (see Akinlabi *et al.*, page 2, sect. 2, 1st paragraph) as would benefit the teachings of Chen *et al.* to include other tonal languages using tonal information.

8. Claims 18 and 19 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chen *et al.* in view of Huang *et al.* as applied to claims 1, 9, and 32 above, and further in view of Chen (2) ("Recognize Tone Languages Using Pitch Information on the Main Vowel of Each Syllable").

As to claims 18 and 19, Chen *et al.* in view of Huang *et al.* teach all of the limitations as in claim 1, above.

However, they do not specifically teach the tonal language comprising Thai and Vietnamese

Furthermore, Chen (2) teaches the tonal language comprising Vietnamese and Thai (see page 4, sect. 7.2, page 4, sect. 7.1).).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech processing system taught by Chen *et al.* and Huang *et al.* with Vietnamese as taught by Chen (2)*et al.*. The motivation to have included such language involves the inclusion of other tone languages such as Vietnamese where tonal information is present (see Chen (2) *et al.*, page 4, and sect. 7.1).

9. Claims 20, 23, and 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chen *et al.* in view of Huang *et al.*

As to claims 20, Chen *et al.* discloses

a speech processing system receiving an input related to one of speech and process the input (see Figure 6, input into microphone 600 , the output of related information would have been obvious to Chen as the system is for use in speech recognition) to provide an output related to one of text and speech processing system (see col. 6, lines 26-36) accessing a module (see col. 3, lines 61-col. 4, lines 8, observations used within the toned phoneme system) derived from a phone set having a plurality of phones for a tonal language (see col. 4, lines 41-44, initials with glides and a second part (final)) comprising a plurality of different tones with different levels of pitch (see col. 4, lines 31-35, each tone has an associated pitch contour), the phones being used to model syllables used in the module (see col. 4, lines 41-44, initials with glides and a second part (final)) and wherein the final part comprises a first portion corresponding to a first

relative pitch and a second portion corresponding to a second relative pitch, wherein the first portion and the second portion jointly and implicitly carry the tonal information (see col. 4, lines 10-13 and col. 4, lines 42-45, the pitch contour varies with time so the pitch changes relative to the portion of the phone i.e. if the phoneme is associated with a rising pitch contour, such a contour is representing a pitch increasing from a base value); and wherein the different levels of pitch comprise at least two categorical levels (see col. 4, lines 33-35, five types of tones), and wherein each portion has a categorical level associated with it (see col. 4, lines 10-15, pitch varies with time and represents a pitch contour. The contour consist of different level or values with respect to time) (E.g. Hence, as the pitch varies over the duration of the syllable the use of categorical levels for each portion vary based on the identified tone. For example, rising tone goes from a low to high value (two categorical levels)).

However, Chen et al. does not specifically disclose the input being text and the output being speech.

Huang *et al.* does disclose the conversion of text to speech from learning methods of model parameters (see Abstract).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech processing system taught by Chen *et al.* to include a text to speech converter as taught by Huang *et al.* The motivation to have included such an element is to have an alternative means for inputting as well as producing a synthesized speech output based upon model

parameters of the system (see Huang *et al.*, Abstract) as would benefit the system of Hon *et al.* by using the tone related information as output speech for producing speech resembling the user.

As to claim 23, Chen *et al.* in view of Huang *et al.* teaches all of the limitations as in claim 20, above.

Furthermore, Chen *et al.* teaches wherein the different levels of pitch comprises five categorical levels (see col. 4, lines 33-35, types of tones), and wherein each portion has a categorical level associated with it (see col. 4, lines 10-15, pitch varies with time and represents a pitch contour. The contour consist of different level or values with respect to time).

As to claim 24, Chen *et al.* in view of Huang *et al.* teaches all of the limitations as in claim 20, above.

Furthermore, Chen *et al.* teaches wherein at least one syllable comprises only the final part having two phones carrying partial tonal information each (see col. 4, lines 14-15 and lines 10-13, lines 31-36, where the second portion comprises one or two phones and the second part contains tone information of the syllable).

10. Claim 22 is rejected under 35 U.S.C. 103(a) as being unpatentable Chen *et al.* in view of Huang as applied to claims 20 above, and further in view of Akinlabi *et al.* (“Tonal Phonology of Yoruba Clitics”).

As to claim 22, Chen *et al.* in view of Huang teaches the phone being associated with a categorical level.

However, they do not specifically disclose the levels of pitch comprising five categorical levels.

Akinlabi *et al.* discloses three tones being associated phonemically (see page 2, sect. 2, lines 1-2).

It would have been obvious to one of ordinary skilled in the art at the time the invention was made to have modified the speech processing system taught by Chen *et al.* in view of Huang. with three categorical levels as taught by Akinlabi *et al.*. The motivation to have included five categorical levels involves the inclusion of other tone languages such as Yoruba, where three tones are present (see page 2, sect. 2, 1st paragraph) as would benefit the teachings of Hon *et al.* to include other tonal languages using tonal information.

Conclusion

11. **THIS ACTION IS MADE FINAL.** Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within

TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the mailing date of this final action.

12. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure.

Yang (US 2001/0010039) is cited to disclose speech recognition of Chinese by using an initial/final similarity vector.

Ao ("A corpus based Mandarin text-to-speech synthesizer") is cited to disclose a TTS system that uses tone and intonation modeling. Cao ("Decision Tree based Mandarin tone model and its application to speech recognition") is cited to disclose decision trees for tonal modeling.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to PARAS SHAH whose telephone number is (571)270-1650. The examiner can normally be reached on MON.-THURS. 7:00a.m.-4:00p.m. EST.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be reached on (571)272-7843. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2626

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/David R Hudspeth/
Supervisory Patent Examiner, Art Unit 2626

/Paras Shah/
Examiner, Art Unit 2626

05/21/2009